

МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
**САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО**

Кафедра теории функций и стохастического анализа

Анализ временных рядов с помощью нейронных сетей

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

Студентки 4 курса 412 группы
направления 01.03.02 – Прикладная математика и информатика
механико-математического факультета
Михеевой Василины Викторовны

Научный руководитель
доцент, к.ф.-м.н. _____ А.В. Шаталина

Зав. кафедрой
зав.каф.,д.ф.-м.н., профессор _____ С.П. Сидоров

Саратов 2025

СОДЕРЖАНИЕ

	Стр.
ВВЕДЕНИЕ	3
1 Основное содержание работы	5
ЗАКЛЮЧЕНИЕ	12
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ.....	13

ВВЕДЕНИЕ

Современные финансовые рынки характеризуются высокой волатильностью и сложностью динамики, что делает задачу прогнозирования экономических временных рядов одной из ключевых проблем в сфере анализа данных и машинного обучения. Традиционные статистические методы, такие как ARIMA, GARCH и линейная регрессия, зачастую оказываются недостаточно эффективными при работе с нелинейными зависимостями и долгосрочными паттернами в данных. В связи с этим возрастает интерес к использованию искусственных нейронных сетей, способных автоматически выявлять сложные закономерности в последовательностях данных.

Среди наиболее перспективных архитектур для анализа временных рядов выделяются рекуррентные нейронные сети (RNN), в частности их модификации — долгая краткосрочная память (LSTM) и управляемый рекуррентный блок (GRU). Эти модели обладают способностью запоминать долгосрочные зависимости, что критически важно для прогнозирования финансовых показателей, где текущие значения могут зависеть от событий, произошедших значительно раньше.

В данной работе исследуется применение моделей LSTM и GRU для прогнозирования ценовых рядов акций крупных компаний, таких как Сбербанк, Роснефть, Tesla, Johnson Johnson, а также криптовалюты Биткоин. Эти активы представляют разные секторы экономики (банковский, нефтегазовый, технологический, фармацевтический и крипто-рынок), что позволяет оценить универсальность и эффективность рассматриваемых методов.

Цель работы — разработка и сравнительный анализ моделей нейронных сетей (в частности, LSTM и GRU) для прогнозирования временных рядов финансовых данных.

Для достижения цели были поставлены следующие задачи:

- Изучить основные понятия финансового рынка
- Рассмотреть принципы работы нейронных сетей

- Описать архитектуры нейронных сетей, применимых к временным рядам
- Провести практический эксперимент
- Сравнить эффективность моделей разных горизонтах.
- Сделать выводы и дать рекомендации по выбору модели в зависимости от типа актива и цели прогнозирования.

Актуальность исследования обусловлена растущим спросом на надежные методы прогнозирования в трейдинге, инвестировании и риск-менеджменте. Результаты работы могут быть полезны как для разработчиков алгоритмических торговых систем, так и для аналитиков, изучающих поведение финансовых рынков с помощью методов машинного обучения.

1 Основное содержание работы

В первом разделе приводятся основные понятия финансового рынка.

Фондовый рынок – это совокупность механизмов и правил, позволяющих осуществлять куплю-продажу акций. В бытовом понимании, фондовый рынок – это место, где происходит торговля ценными бумагами.

Основные виды ценных бумаг

- Акция
- Облигация

Фундаментальный анализ – это детальное, разбитое на отдельные фундаментальные показатели, исследование текущего состояния и экономических перспектив компании (сектора, региона, страны). При этом основное внимание, как правило, уделяется анализу текущих и прогнозу будущих доходов, процентным ставкам, а также оценке потенциальных преимуществ и рисков.

Технический анализ – это система правил, по которым интерпретируются рыночные движения и которые приводят к определенным выводам относительно возможных движений в будущем.

Во втором разделе приводится определение абстрактного нейрона.

Определение 1.1. Абстрактный нейрон – это четверка (x, w, φ, y) , где $x^T = (x_0, x_1, \dots, x_n)$ – входной вектор, $w^T = (w_0, w_1, \dots, w_n)$ – вектор весов, где $x_0 = 1$ и $w_0 = b$ – смещение, а φ – функция активации, которая определяет результирующую функцию $y = \varphi(x^T, w) = \varphi(\sum_{i=0}^n w_i x_i)$.

В третьем разделе описана архитектура нейронных сетей.

Нейронная сеть (искусственная нейронная сеть, ИНС) – это математическая и вычислительная модель, вдохновлённая биологическими нейронными сетями (например, мозгом). Она состоит из взаимосвязанных искусственных нейронов (узлов), обрабатывающих информацию по принципам, схожим с работой нервной системы живых организмов.

Ключевые компоненты нейронной сети:

Нейроны (узлы) – базовые единицы, которые принимают входные данные, выполняют вычисления и передают результат.

Связи (веса) – параметры, определяющие силу влияния одного нейрона на другой.

Функция активации – нелинейное преобразование (например, ReLU, сигмоида), добавляющее сложность модели.

Слои – группы нейронов:

Входной слой – получает исходные данные.

Скрытые слои – промежуточные вычисления.

Выходной слой – итоговый результат (например, класс объекта или числовой прогноз).

Как работает нейронная сеть:

Прямое распространение (Feedforward) – данные проходят через слои, преобразуясь с помощью весов и функций активации.

Обучение (Backpropagation) – на основе ошибки предсказания корректируются веса (обычно через градиентный спуск).

Основные виды нейронных сетей

1. По структуре (архитектуре)

- Искусственные нейронные сети (ИНС, ANN) – базовый тип, состоящий из входного, скрытого и выходного слоёв.
- Многослойный перцептрон (MLP, Multilayer Perceptron) – классическая сеть с полносвязными слоями.
- Свёрточные нейронные сети (CNN, Convolutional Neural Network) – используются для обработки изображений, видео и других пространственных данных.
- Рекуррентные нейронные сети (RNN, Recurrent Neural Network) – работают с последовательностями (текст, временные ряды).
- Долгая краткосрочная память (LSTM, Long Short-Term Memory) – улучшенная RNN для долгосрочных зависимостей.

- Управляемые рекуррентные блоки (GRU, Gated Recurrent Unit) – упрощённая версия LSTM.
- Автокодировщики (Autoencoders) – сети для сжатия и восстановления данных.
- Генеративно-сопоставительные сети (GAN, Generative Adversarial Network) – состоят из генератора и дискриминатора, создают новые данные (например, изображения).
- Трансформеры (Transformers) – архитектура на основе механизма внимания (attention), популярна в NLP (BERT, GPT).
- Сети с вниманием (Attention Networks) – используют механизм внимания для выделения важных частей данных.
- Остаточные сети (ResNet, Residual Network) – содержат "пропускающие" соединения для обучения глубоких сетей.
- Графовые нейронные сети (GNN, Graph Neural Network) – работают с графами (социальные сети, молекулы).

2. По типу обучения

- С учителем (Supervised Learning) – обучение на размеченных данных (классификация, регрессия).
- Без учителя (Unsupervised Learning) – поиск закономерностей без меток (кластеризация, автоэнкодеры).
- С частичным привлечением учителя (Semi-supervised Learning) – комбинация размеченных и неразмеченных данных.
- С подкреплением (Reinforcement Learning) – обучение через взаимодействие со средой (DQN, AlphaGo).

В четвертом разделе изучается метод обучения НС - метод обратного распространения ошибки.

Метод обратного распространения ошибки — это алгоритм обучения нейронных сетей, который вычисляет градиент ошибки по весам сети с помощью цепного правила дифференцирования (chain rule). Он позволяет эффективно обновлять веса сети, минимизируя функцию потерь.

Основная идея метода обратного распространения ошибки

Нейронная сеть делает предсказание (прямой проход, forward pass), сравнивает его с истинным значением (через функцию потерь), а затем распространяет ошибку назад (backward pass), корректируя веса для уменьшения ошибки.

Пятый раздел посвящен построению моделей рекуррентных нейронных сетей - LSTM и GRU.

Используются библиотеки:

NumPy, Pandas — работа с массивами и таблицами.

MinMaxScaler — нормализация данных.

Keras (LSTM, GRU, Dropout, BatchNormalization) — архитектура моделей.

yfinance — загрузка исторических котировок акций и криптовалют.

os, matplotlib — файловая система и построение графиков.

LSTM — это особый тип рекуррентной нейронной сети (RNN), разработанный для решения проблем долгосрочных зависимостей и затухающих градиентов, присущих обычным RNN.

Основная идея LSTM

LSTM вводит механизм "вентилей"(gates), которые регулируют поток информации:

- Забывание ненужных данных,
- Запоминание важной информации,
- Обновление состояния ячейки.

Это позволяет сети эффективно работать с длинными последовательностями (например, текстами, временными рядами).

GRU (Gated Recurrent Unit) — это упрощённая версия LSTM, разработанная для более эффективного обучения рекуррентных сетей с сохранением способности улавливать долгосрочные зависимости.

Основная идея GRU

GRU объединяет входной вентиль (input gate) и вентиль забывания (forget gate) из LSTM в один "вентиль обновления"(update gate), а также

заменяет ячейку состояния (cell state) и скрытое состояние (hidden state) на единый механизм обновления.

В коде используются эти два типа рекуррентных нейронных сетей:

1. LSTM (Long Short-Term Memory)

Архитектура:

- Первый LSTM-слой с 64 нейронами и `return_sequences=True`, что означает, что он возвращает выход для каждого временного шага (необходим для передачи в следующий рекуррентный слой).
 - Dropout-слой (20%) — для регуляризации, предотвращает переобучение.
 - Второй LSTM-слой с 32 нейронами (`units // 2`).
 - BatchNormalization (опционально) — для стабилизации и ускорения обучения.
 - Финальный Dense-слой с одним выходом — предсказание следующей цены.
- Используется регуляризация (L1, L2 или L1L2),.

2. GRU (Gated Recurrent Unit)

Архитектура аналогична LSTM, только вместо LSTM-слоев используются GRU-слои.

Отличие GRU: чуть проще и быстрее обучается по сравнению с LSTM, с меньшим количеством параметров.

Цель: предсказывать следующую цену закрытия криптовалюты или акции на основе предыдущих 60 временных шагов, включающих цену закрытия и скользящую среднюю (SMA).

Обе модели компилируются с оптимизатором Adam и функцией потерь MSE (среднеквадратичная ошибка).

Загрузка и подготовка данных

Скачивает данные с Yahoo Finance.

Вычисляет скользящую среднюю на 20 дней (SMA20).

Масштабирует два признака: цену закрытия и SMA20 с помощью MinMaxScaler.

Создание выборки

Создает обучающие примеры: каждый пример состоит из 60 дней истории (двумерный вход: Close + SMA).

Целевое значение — Close на +1, +15, +30 дней вперёд (forecast_horizon=1).

Считаются метрики:

MSE — среднеквадратичная ошибка,

RMSE — её корень,

RMSE % — RMSE в процентах от средней цены,

MAPE — средняя абсолютная процентная ошибка.

Определяется лучшая модель.

Были взяты данные 12 активов:

SOL-USD – Криптовалюта Solana, торгуемая к доллару США. Популярный блокчейн-проект, ориентированный на высокую скорость и масштабируемость.

DOGE-USD – Dogecoin к доллару. Мем-криптовалюта, изначально созданная как шутка, но ставшая популярной благодаря сообществу и поддержке Илона Маска.

ETH-USD – Ethereum к доллару. Вторая по капитализации криптовалюта в мире, основа для большинства смарт-контрактов и децентрализованных приложений.

SBER.ME – Акции Сбербанка России (тикер на Московской бирже: SBER). Один из крупнейших банков Восточной Европы.

TSLA – Акции Tesla Inc. (США). Производитель электромобилей и инновационных энергетических решений во главе с Илоном Маском.

BTC-USD – Биткоин к доллару. Первая и самая известная криптовалюта, цифровое золото.

TCSG.ME – Акции TCS Group Holding (владелец Тинькофф Банка). Один из лидеров финтеха в России.

JNJ – Акции Johnson Johnson. Многонациональная корпорация, производитель медицинских товаров, фармацевтики и потребительской продукции.

PG – Акции Procter Gamble. Один из крупнейших в мире производителей товаров повседневного спроса (бренды: Ariel, Pampers, Gillette и др.).

GAZP.ME – Акции Газпрома (на Московской бирже). Крупнейшая российская энергетическая компания, экспортирующая природный газ.

XRT – Биржевой фонд (ETF) SPDR SP Retail ETF. Отслеживает акции розничных компаний США.

VLO – Акции Valero Energy Corporation. Американская нефтеперерабатывающая компания, занимающаяся производством и маркетингом топлива и нефтехимии.

Был проведен прогноз соответственно на 1, 15 и 30 дней.

GRU в целом демонстрирует более стабильную и точную работу на коротких и средних горизонтах, особенно при работе с волатильными активами. LSTM начинает обгонять её только на длинных горизонтах прогнозирования, где важно учитывать дальние зависимости.

Выбор модели зависит от цели прогноза:

- Для краткосрочной торговли – GRU.
- Для долгосрочного планирования – LSTM.

В приложении представлены исходные программные коды реализации.

ЗАКЛЮЧЕНИЕ

В данной работе был проведён анализ акций и криптовалют с использованием моделей глубокого обучения — LSTM и RNN. Эти архитектуры были выбраны благодаря их способности эффективно обрабатывать последовательные данные и выявлять сложные временные зависимости в финансовых временных рядах.

В ходе эксперимента были созданы 2 хороших нейронных сети и выявлены особенности использования LSTM и GRU. Нейросетевые модели, GRU и LSTM, доказали свою эффективность в прогнозировании финансовых временных рядов, но их успешное применение на практике требует учёта специфики рыночных данных и постоянного совершенствования алгоритмов.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- 1 Галушкин, А. И. Нейронные сети: основы теории / А. И. Галушкин. М.: ГЛТ, 2012. - 496 с.
- 2 Классификация и типы нейронных сетей [Электронный ресурс]: [сайт]. URL: <https://pythonru.com/primery>-(дата обращения: 31.02.24). Загл. с экрана. Яз. рус.
- 3 Что такое НС [Электронный ресурс]: [сайт]. URL: <https://habr.com/309508/>(дата обращения 20.03.24). Загл. с экрана.-Яз. рус.
- 4 Гафаров, Ф. М. Искусственные нейронные сети и приложения / Ф. М. Гафаров - Казань: Изд во Казан. ун-та 2018.-121 с.
- 5 Основы нейронных сетей [Электронный ресурс]: [сайт]. URL <https://www.dmitrymakarov.ru/intro/neural-21/> (дата обращения 02.03.24). - Загл. с экрана. Яз. рус
- 6 Твардовский В.В. "Теория и практика торговли на фондовом рынке"/ Твардовский В.В. изд. третье, испр. - СмартБук, 2011.—296 с.
- 7 Ovidiu Calin "Deep Learning Architectures A Mathematical Approach"/ Ovidiu Calin изд. Springer, 2021. -768 с.
- 8 Что такое искусственные нейронные сети? [Электронный ресурс]: [сайт]. URL <https://cloud.ru/ru/blog/neural-networks> (дата обращения 02.04.24). - Загл. с экрана. Яз. рус
- 9 Фундаментальный и технический анализ [Электронный ресурс]: [сайт]. <https://gog.su/FpzN> (дата обращения 09.03.24). - Загл. с экрана. Яз. рус
- 10 Метод обратного распространения ошибки [Электронный ресурс]: [сайт]. <https://education.yandex.ru/handbook/ml/article/metod-obratnogo-rasprostraneniya-oshibki> (дата обращения 18.03.24). - Загл. с экрана. Яз. рус